

Eur. J. Clin. Chem. Clin. Biochem.
Vol. 31, 1993, pp. 233–237

© 1993 Walter de Gruyter & Co.
Berlin · New York

Serum Iron and Proficiency Testing: A Dilemma

By B. G. Blijenberg, H. A. Roetering and H. J. Brouwer

Academic Hospital Rotterdam-Dijkzigt, Department of Clinical Chemistry, Rotterdam, The Netherlands

(Received November 5, 1992/January 8, 1993)

Summary: We present and discuss observations regarding the results of the serum iron determination as seen in various surveys. The reference method proved to score different in our hands judging on the performance in the Dutch and German quality assessment schemes. Furthermore, it is clear and frustrating that these scores, representing the most accurate results, are becoming more and more labeled “out of acceptable range”.

Introduction

In the last two decades numerous reports on the analytical aspects of the determination of serum iron have been published. These concern mainly the simplification of many procedures, the development and application of new chromogens and the adaptations to various instruments. A step forward, unfortunately not always used in the many reported comparisons, was the introduction of the reference method by the International Committee for Standardization in Haematology (ICSH) in 1978 (1). A rough consensus of several reports indicates that values obtained with this method are elevated by 0–10%. This is illustrated by the three recent reports presented in table 1. The differences are assumed to be related to the removal of iron from transferrin, the release of iron from haemoglobin, the precipitation of hydrolysed ferric hydroxyde and the chelation of copper and zinc (2–4).

In the early eighties we decided to introduce the reference method in our laboratory for routine purposes. Participation in the quality control programmes (5, 6) of the Dutch Quality Assessment Foundation (SKZL) followed, and to our surprise they revealed greater differences between our results and the consensus values than expected. Figure 1 shows the results of the years 1986 to 1990. We tried to find an explanation for this phenomenon. The use of various standard preparations, commercial and home-made (iron wire and ferrous salts) did not re-

Tab. 1. Comparison of the results for serum iron in patient samples as obtained with the reference method (x) and a method without protein precipitation (y).

Authors	Regression equation	n	r
Johnson et al. (10)	$y = 1.026x - 1.47$	28	0.985
Makino et al. (11)	$y = 1.00x - 1.43$	70	0.995
Valcour et al. (12)	$y = 0.947x + 1.43$	200	0.982

n = number of samples

r = coefficient of correlation

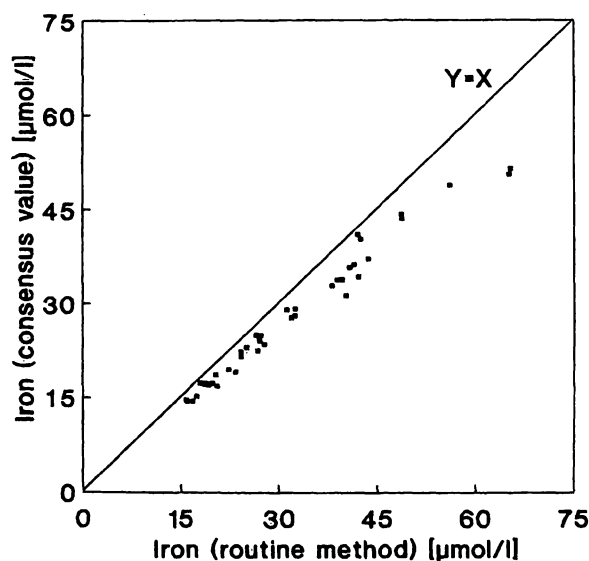


Fig. 1. The results of 42 survey samples over the years 1986–1990 (Dutch Quality Assessment Foundation).

Statistics:

y (= consensus) = $0.86x$ (= routine) + 0.7 ; $r = 0.99$

move the discrepancies. The results of the analyses with patient samples as well as commercial samples remained the same using the different standards. Only the values for the horse serum-based quality control samples from the Dutch National Institute of Public Health were in agreement with our results. These samples were in use for many years in the majority of Dutch clinical chemistry laboratories. For iron, a peer group consisting of 6–8 laboratories provided the various iron results by applying the same reference method as described above (1). To our knowledge, during those years, these serum samples were the only samples commercially available with reference method assigned values. During that period, we always obtained values within 0.5 $\mu\text{mol/l}$ for our determinations of recovery, using various preparations with reference method-based assigned values (low concentrations ranging from 20 to 25 $\mu\text{mol/l}$ and high concentrations from 35 to 40 $\mu\text{mol/l}$).

An example of our satisfaction with the performance of our routine method is given by the results obtained with the last lots of these control samples. Over a period of 39 days we found an average value of 23.9 $\mu\text{mol/l}$ with a standard deviation of 1.1 $\mu\text{mol/l}$ (stated true value $23.7 \pm 1.0 \mu\text{mol/l}$) for the low concentration, and 39.8 $\mu\text{mol/l}$ (S.D. 1.4 $\mu\text{mol/l}$) (stated true value $40.1 \pm 1.0 \mu\text{mol/l}$) for the high concentration.

In 1990 the production of these sera unfortunately stopped. For various reasons, including the assessment of our serum iron determination, we decided in

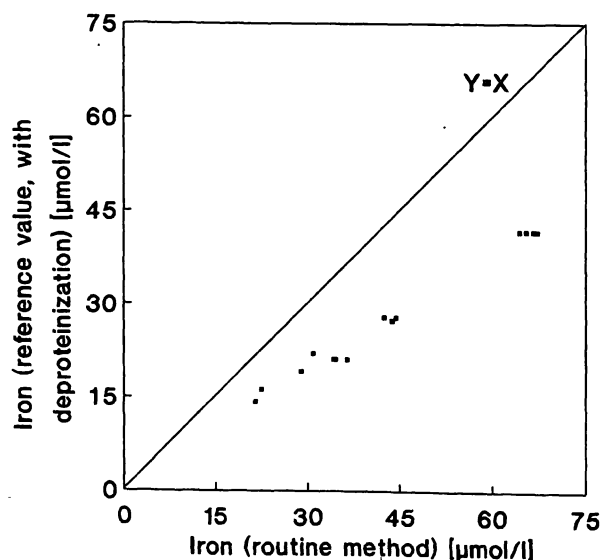


Fig. 2. The results of 14 survey samples over the years 1990–1992 (German Quality Assessment Scheme).

Statistics:

x = routine method and y = reference value with deproteinization

$y = 0.63x + 0.1$; $r = 0.99$

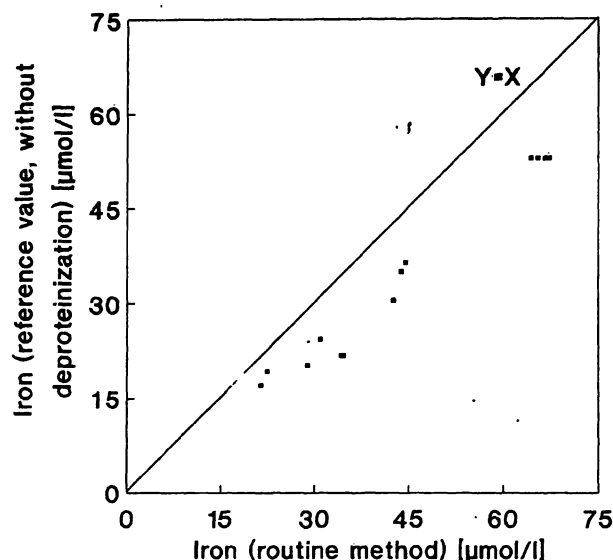


Fig. 3. The results of 14 survey samples over the years 1990–1992 (same samples as depicted in fig. 2).

Statistics:

x = routine method and y = reference value without deproteinization

$y = 0.83x - 2.1$; $r = 0.98$

1989 to participate in the control scheme of the German Society for Clinical Chemistry. The results were surprising. These are shown for the years 1990, 1991 and 1992 in figures 2 and 3, in which a method involving deproteinization is compared with a non-deproteinization method. A more thorough study was therefore needed, and this is presented below.

Materials and Methods

Materials

All chemicals were obtained from Merck (Darmstadt, Germany).

Patient serum samples were chosen at random from various departments in our hospital.

Quality control specimens were from the Dutch Quality Assessment Foundation (Dr. H. Baadenhuijsen, Nijmegen) and the German Reference Institution (Prof. Dr. G. Röhle, Bonn) and were used in various surveys.

Methods

Our method for routine serum iron is the same as that published as a reference method (1).

The atomic absorption spectrometry (AAS) method for iron represents a conventional application of the instrument. Briefly, serum is diluted 100-fold with demineralized water, ashed at 1100 °C and atomized at 2100 °C in a furnace and measured at 248.3 nm.

The commercial test, iron without deproteinization, was obtained from Boehringer Mannheim GmbH (Germany), catalogue number 759422.

Instrumentation

All colorimetric measurements were done with a Philips PU 8700 ultraviolet/visible spectrophotometer and the AAS measurements with a Philips PU 9200 X atomic absorption spectrometer.

Statistics

For statistical evaluation we used the method of *Passing & Bablok* (9).

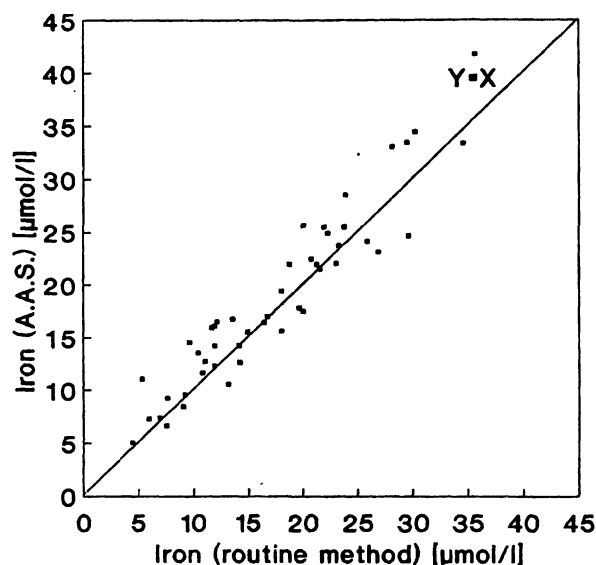


Fig. 4. The results of 47 patient samples chosen at random (non-lipaemic, non-icteric and non-haemolytic).

Statistics:

$$y (= \text{AAS}) = 0.96x (= \text{routine}) - 0.4; r = 0.95$$

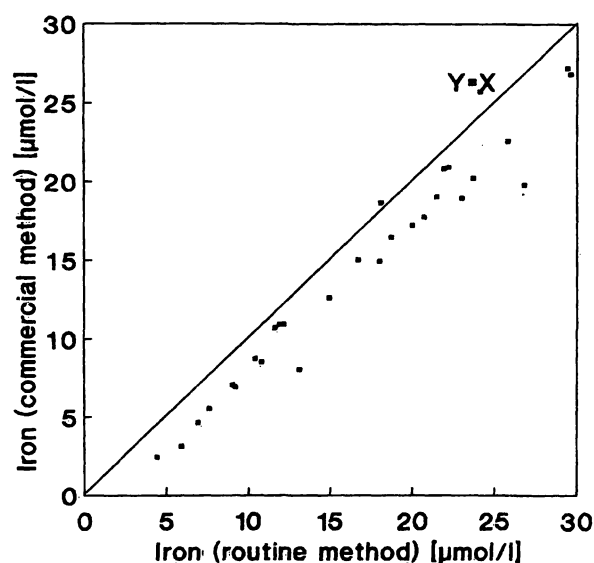


Fig. 5. The results of 29 patient samples (part of the set mentioned in fig. 4).

Statistics:

$$y (= \text{commercial}) = 0.98x (= \text{routine}) - 2.0; r = 0.98$$

Results

As comparison methods we chose a procedure based on atomic absorption spectrometry and a colorimetric method without protein precipitation (see Materials and Methods). The graph in figure 4 compares the results obtained with the reference method and the atomic absorption spectrometric method, using non-icteric and non-haemolytic patient samples. Part of this set of patient samples was also analysed with the colorimetric method (see fig. 5).

Finally, we analysed 21 commercial quality control samples with both colorimetric methods i.e. the precipitation and the non-precipitation method (see fig. 6).

Discussion

From our own internal quality control system, we knew that we had routinely applied a stable and precise iron determination over the years. The procedure was time consuming compared with automated methods but satisfying. Its accuracy has often been a matter of debate in relation to the results of surveys. At first, the control samples of the Dutch Institute for Public Health proved to be of great value, and the results from these reassured us (see Introduction). Later, this feeling was supported by the results obtained with atomic absorption spectrometry (see fig. 4).

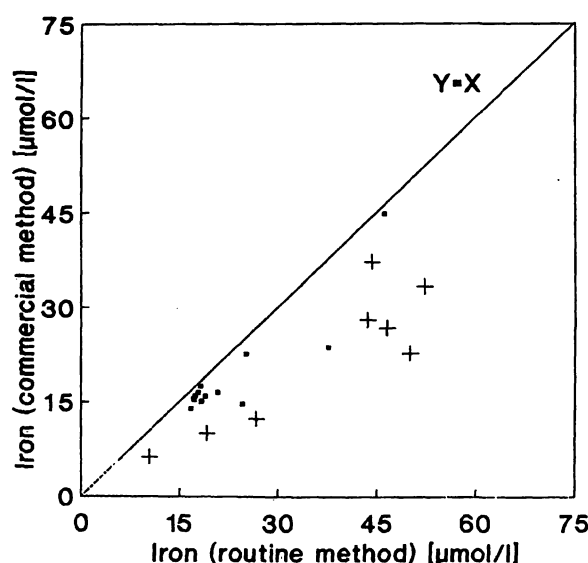


Fig. 6. The results of 21 survey samples.

Statistics:

$$y (= \text{commercial}) = 0.72x (= \text{routine}) + 1.8; r = 0.84$$

The values marked + refer to icteric and/or turbid specimens.

Nevertheless, at the end of the eighties, we became more and more isolated in the scoring system of the External Dutch Quality Control System. Knowing that we used the best available method in terms of accuracy, and also knowing that our precision was excellent, we were still "punished", the majority of participants using methods without deproteinization. Participation in the German scheme made this matter even worse, as can be seen in figures 2 and 3.

Our unsatisfactory feelings were also supported by the results depicted in figure 6. In our opinion, the rather large scatter is due to the matrix of the commercial samples. Eight of the 21 samples were icteric and/or turbid, and these factors generally influence the results of non-precipitation methods. It is difficult to determine the effect of such influences in survey results.

It was encouraging to read the recent article of *Tietz et al.* entitled: "Are clinical laboratory proficiency tests as good as they can be*" (7). It is beyond the scope of our article to discuss this matter here, but it also describes, for example, the unsatisfactory relationship between the determination of serum iron and proficiency testing.

All our practical work was finished when the article of *Tietz et al.* appeared. Since *Tietz et al.* do not present practical evidence or literature references, it is not easy to perform a quantitative comparison of our respective experiences. Nevertheless, we fully agree with their plea for the use of reference methods, which is in accordance with the experience of quality control in Germany (8).

It is not clear why the results of the German Quality Assessment Scheme differ so much from the Dutch (figs. 1, 2 and 3). We assume that it is related to the method of comparison. In The Netherlands a consensus value is used for performance evaluation, while in Germany peer comparison is applied for serum iron. In the context of the German scheme we only used two sets of data: results obtained with a precipitation method and those with a non-precipitation method.

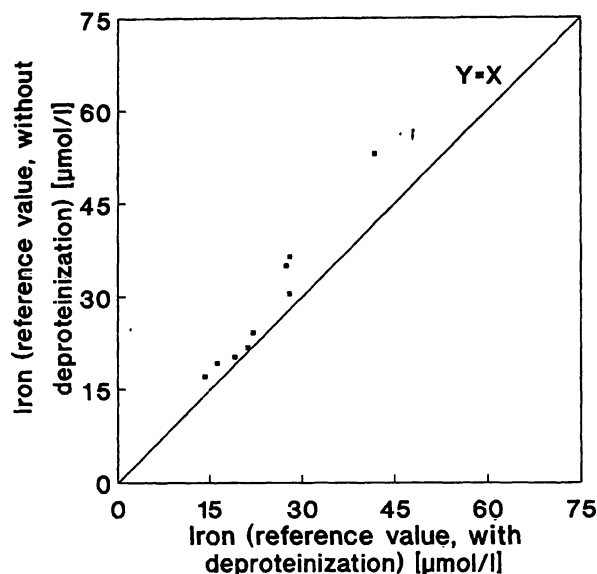


Fig. 7. The results of the same survey samples as depicted in figures 2 and 3.

N. B. 9 points are represented instead of 14 because some samples were used in 2 surveys.

It is surprising that the precipitation method always gives lower results than the non-precipitation method in the peer group (see fig. 7). It is also surprising that the results obtained with the precipitation method used by the peer group differ so much from ours. Ultimately, the reference method is also a precipitation method.

In summary, we have at the moment no practical solution for this dilemma. We still feel that our analytical performance is good. For the future, we advocate a clear reference system for serum iron, supported by samples with reference method-based iron values. This means that choices are necessary. The first choice has already been made: a reference method does exist. The question arises as to how the method principle can be applied for both precipitation and non-precipitation methods. This depends on the development of well characterized reference samples, and it may mean that these samples must have a human serum matrix. In this way it should be possible to assess methods and thus circumvent problems of the kind described above.

References

1. International Committee for Standardization in Haematology (1978) Recommendations for measurement of serum iron in human blood. *Brit. J. Haematol.* **38**, 291–294.
2. Itano, M. (1978) CAP comprehensive chemistry: Serum iron survey. *Amer. J. Clin. Pathol.* **70**, 516–522.
3. Cavill, I. (1982) Diagnostic methods. *Clin. Haematol.* **11**, 259–273.
4. Derman, D. P., Green, A., Bothwell, T. H., Graham, B., Mc Namara, L., Mac Phail, A. P. & Baynes, R. D. (1989) A systematic evaluation of bathophenanthroline, ferrozine and ferene in an ICSH-based method for the measurement of serum iron. *Ann. Clin. Biochem.* **26**, 144–147.
5. Jansen, A. P., Van Kampen, E. J., Leijnse, B., Meyers, C. A. M. & Van Munster, P. J. J. (1977) Experience in The Netherlands with an external quality control and scoring system for clinical chemistry laboratories. *Clin. Chim. Acta* **74**, 191–201.
6. Jansen, R. T. P. & Jansen, A. P. (1980) A coupled external/internal quality control program for clinical laboratories in The Netherlands. *Clin. Chim. Acta* **107**, 185–201.
7. Tietz, N. W., Rodgerson, D. O. & Laessing, R. H. (1992) Are clinical laboratory proficiency tests as good as they can be? *Clin. Chem.* **38**, 473–475.
8. Stamm, D. (1988) Qualitätssicherung der quantitativen Bestimmungen im Laboratorium. Neue Richtlinien der Bundesärztekammer. *Dtsch. Ärztebl.* **85**, B-517–B-532.
9. Passing, H. & Bablok, W. (1983) A new biomedical procedure for testing the equality of measurements from two different analytical methods. *J. Clin. Chem. Clin. Biochem.* **21**, 709–720.
10. Johnson, D. J. & Williams, H. L. (1990) Rapid spectrophotometric assay of serum iron. *Clin. Chim. Acta* **189**, 199–204.
11. Makino, T., Kiyonaga, M. & Kina, K. (1988) A sensitive, direct colorimetric assay of serum iron using the chromogen, nitro-PAPS. *Clin. Chim. Acta* **171**, 19–28.
12. Valcour, A. A., Krzymowski, G., Onoroski, M., Bowers, G. N. & McComb, R. B. (1990) Proposed reference method for iron in serum used to evaluate two automated iron methods. *Clin. Chem.* **36**, 1789–1792.

Dr. B. G. Blijenberg
Academic Hospital Rotterdam-Dijkzigt
Department of Clinical Chemistry
Dr. Molewaterplein 40
NL-3015 GD Rotterdam
The Netherlands

